1C

УДК 681.3:519.9

В.В. Сторож
Донецкий национальный технический университет, Украина
Украина, 83050, г. Донецк, пр. Богдана Хмельницкого, 84

Принципы реализации семантики в естественных и искусственных системах

V.V. Storozh
Donetsk National Technical University, Ukraine
Ukraine, 83050, c. Donetsk, Bogdana Khmelnitskogo av.84

Principles of Realization of Semantics in Natural and Artificial Systems

В.В. Сторож
Донецький національний технічний університет, Україна
Україна, 83050, м. Донецьк, пр. Богдана Хмельницького, 84

Принципи реалізації семантики в природних та штучних системах

Семантика сообщения формируется как собственно сообщением, так и воспринимающей это сообщение интеллектуальной системой. Человек в процессе понимания сообщения формирует на его основе некоторые интеллектуальные целостности, – гештальты, идеи, схемы и т.п. Эти выделенные схемы в той или иной мере соответствуют внутренним схемам, основанным на желаниях и запросах.
Ключевые слова: смысл и значение, гештальт, схема, интеллектуальная система.

Semantics of the message is formed as the message proper, and intellectual system perceiving this message. The person in the course of understanding of the message forms on its basis some intellectual integrity, - a Gestalt, ideas, schemes, etc. These allocated schemes to some extent correspond to the internal schemes based on desires and inquiries.
Key words: sense and meaning, gestalt, schema, intellectual system.

Семантика повідомлення формується як власне повідомленням, так і сприймаючій це повідомлення інтелектуальною системою. Людина в процесі розуміння повідомлення формує на його підставі деякі інтелектуальні цілісності, – гештальти, ідеї, схеми і т.п. Ці виділені схеми в тій чи іншій мірі відповідають внутрішнім схемам, що засновані на бажаннях та запитах.
Ключові слова: смисл та значення, гештальт, схема, інтелектуальна система.

Введение

Компьютерная обработка текстовой и речевой информации на уровне смысла, значения, понимания информации весьма примитивна, и ни в какой мере не может конкурировать с человеком.

Так, несмотря на распространенность компьютерных переводчиков, они могут осуществлять лишь пословный перевод (и то во многих случаях очень неточный), зачастую неспособны передать смысл уже на уровне предложения, и совершенно не понимают смысла текста в целом.

При работе в Интернете и вообще с библиотеками информации часто необходимо провести поиск информации, скачать, проанализировать полученную информацию и

в результате получить «выжимку» на 1 – 3 страницы. Как правило, на любой вопрос мы получаем количество ссылок от тысяч до десятков миллионов. Уже здесь возникают проблемы избытка информации по выбору нужных нам ссылок. Но пусть мы просмотрели значительную часть ссылок, отсеяли рекламу, дилетантов и т.д., скачали сами тексты. Это могут быть тысячи статей и сотни и тысячи книг на не очень понятном английском языке. Анализ всей этой литературы может занять годы. А требуются не более нескольких часов, в крайнем случае, дней. Хорошо бы, если бы все указанное сделал компьютер, но он пока на это не способен.

Представляется, что следующий этап развития компьютерной семантики должен быть направлен именно на решение указанных выше задач и сравнимых с ними по трудности. Для этого необходимо решение целого ряда принципиальных вопросов:

1. Понять, чем отличаются процессы построения семантики (смысла, значения и т.п.) в естественных и искусственных системах;
2. Понять хотя бы на общем уровне, каким образом, с помощью каких механизмов, инструментов и процессов формируется семантика в естественных и искусственных системах;
3. Понять, каким образом можно наиболее адекватно моделировать процессы построения семантики в естественных системах, а затем реализовывать эти модельные представления в искусственных системах.

Поиску ответов на указанные вопросы и посвящена данная работа.

1. Проблемы получения значения сообщения

Используем в данной работе по отношению к тексту и речи обобщающий термин «сообщение». В широком смысле под сообщением можно понимать любую информацию: текст, речь, схемы, формулы, и даже произведения искусства. В данной работе будем рассматривать только те типы сообщений, которые либо непосредственно выражаются с помощью понятий (текст, речь), либо в значительной мере выразимы с помощью понятий (схемы, формулы и т.п.). Такое ограничение вызвано тем, что два указанных типа сообщений существенно различны. Смысл понятийных и частично формализованных сообщений понимается субъектами более или менее одинаково, определяется в основном рассматриваемой предметной областью, и при необходимости может быть сформулирован более строго. В то же время произведения искусства выражаются (за исключением литературы) без использования понятий (например, музыка и живопись), трактуются различными субъектами существенно различным образом (т.е. имеют существенно различный смысл), и требуют для своего обоснования сущностного понимания таких понятий как сознание, душа, смысл жизни, наслаждение, романтика и т.д. Требуемый для такого восприятия уровень семантики находится, по-видимому, за горизонтом возможностей компьютерных систем современного типа.

В работах представителей различных дисциплин, – лингвистики [1], когнитивной лингвистики [2], искусственного интеллекта [3-5] делаются выводы о том, что понимаемое человеком содержание текста и речи, называемое также смыслом, значением или семантикой, значительно богаче, чем это представляется в формализованных компьютерных моделях.

В самом общем виде схема получения значения сообщения приведена на рис. 1. В компьютере семантика текста сводится к сумме семантики отдельных слов текста и их связей. Человек же в процессе интерпретации и понимания текста создает существенно отличающееся от исходного теста содержание, в результате чего может по-

явиться ряд принципиально новых понятий, важных для понимания текста и отсутствующих в исходном варианте, новые связи, выводы и т.д. У компьютера «смысл» есть содержание собственно текста. База данных при этом относительно невелика и состоит из краткого описания предметной области и правил связывания единиц текста. У человека извлечение смысла из текста направляется его желаниями и поддерживается обширной базой знаний (всем жизненным опытом человека). В сущности, из рис. 1 можно «вытащить» целый ряд более детализированных схем для человека и компьютера.

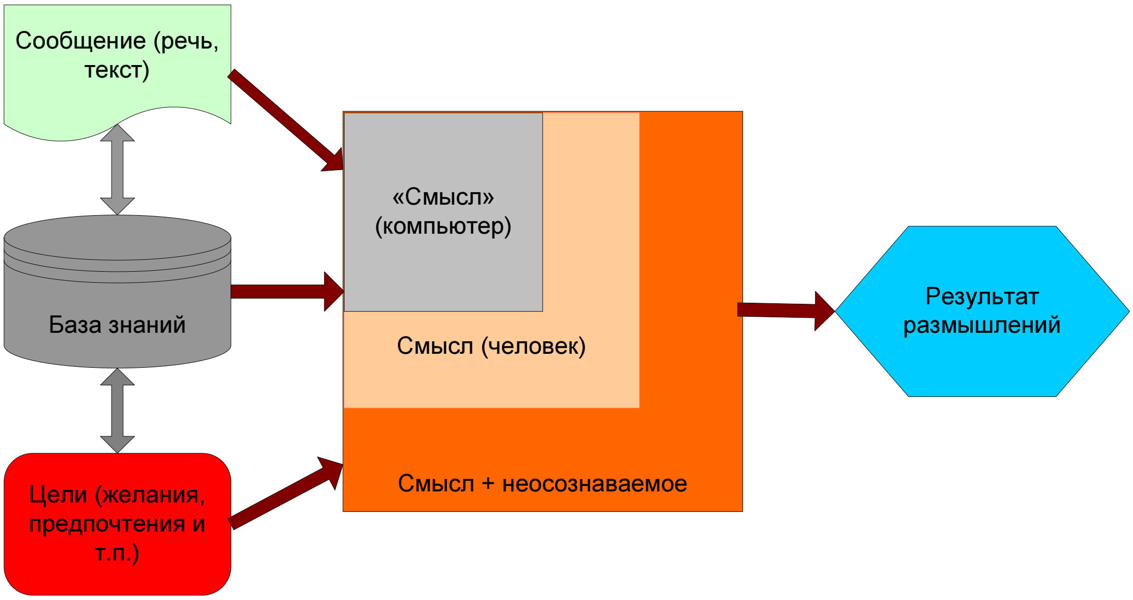


Рисунок 1 – Общая схема получения значения

Для разработки следующего уровня компьютерной семантики мало просто тезиса о различии объемов смысла у человека и компьютера. Необходимо понять источники этого различия и с их учетом найти инструменты для более адекватного представления семантики в компьютерных системах. Для решения этой задачи проанализируем хорошо известный в логике и когнитивной лингвистике термин «объем понятия», требования к объему понятия с точки зрения представления реальности, способы изменения объема понятий и т.п.

Под объемом понятия в данной работе понимается, с одной стороны, величина области реальности (действительно существующей, или воображаемой), представляемая понятием, а с другой, – объем содержания как величина и разнообразие описания.

Вопрос об оптимальном отображении реальности с помощью понятий в лингвистике и компьютерной семантике, как правило, не рассматривается, а между тем он исключительно важен. В подтверждение этой точки зрения приведем примеры наличия методов решения сложных задач, в которых решение ищется в виде некоторых оптимальных по сложности информационных объектов, в какой-то мере аналогичных понятиям. В частности, это метод группового учета аргументов (МГУА), в котором решение ищется в виде набора полиномов определенной сложности и вида, методы генетических алгоритмов, в которых решение ищется с помощью модифицированных до оптимального уровня хромосом и т.д. И еще пример непосредственно из области языков. В русском языке имеются три-четыре понятия, описывающих

снег (наст, пороша), а в эскимосском – около 50-ти. Эскимосы выделяют как отдельный элемент реальности и называют, в частности, снег, который олени могут пробить копытами и кормиться, снег, по которому могут проехать нарты и др. Вследствие этого у них не возникает проблем ни с передачей информации относительно снега, ни с ее эффективным использованием. А если бы эскимосы использовали русское понятие снега, то они не смогли бы различить снег, который могут разгрести олени, снег, в котором не утонут нарты и т.п., и, скорее всего, погибли бы. Из этого и аналогичных примеров можно сделать вывод, что и в человеческом языке существует принцип оптимального по объему и содержанию понятийного представления интересующей предметной области. Подкрепляя эти соображения аналогией с методами МГУА и генетическими алгоритмами можно утверждать, что для различных предметных областей и задач оптимальный объем и содержание одних и тех же понятий, вообще говоря, различны.

Далее рассмотрим некоторые наиболее известные причины и методы «подгонки» содержания понятий под требования наиболее точного и оптимального представления текущей задачи.

1. Чрезвычайно широко используемым источником получения новых значений являются сочетания различных частей речи, – например, существительного и прилагательного, и вообще различного рода грамматические формы, – дополнения, именные сказуемые и др. В качестве примера рассмотрим содержание прилагательного «жесткая» в двух выражениях: жесткая дискуссия и жесткая стыковка. Для указанных случаев в слове «жесткая» можно выделить как одинаковое, так и различающееся содержание. Одинаковое – это указание на некоторую чрезвычайность и превышение средних тенденций при реализации процесса. Различающееся – это способы реализации чрезвычайности. Для стыковки это физический толчок или удар, для дискуссии – это психологическое противостояние, ультимативные заявления, неприятие противоположной точки зрения и т.д. Отметим, что в рассматриваемых случаях содержание слова «жесткий» раскрывается (детализируется, уточняется и т.п.) существенно различным образом как на понятийном, так и на схемном уровнях. При этом суть взаимодействия двух понятий сводится не просто к суммированию их содержаний (хотя и этот процесс в некоторой степени происходит), а к генерации принципиально нового, не содержащегося в исходных понятиях содержания! Это тот вид интеллектуальной деятельности, с которой превосходно справляется человек, и совершенно не справляются современные компьютерные программы.

2. Еще один чрезвычайно распространенный метод получения новых значений, это использование фигур речи (тропы), к наиболее распространенным и изученным из которых относятся аналогии и метафоры. В работах Дж. Лакоффа и его последователей [6], [7] детально рассмотрено влияние метафор на процессы человеческого мышления и генерацию нового содержания. Лакофф ввел представление о структурной метафоре [6, с. 35]: это случаи, когда содержание одного концепта (понятия) структурируется на частичном (как правило) базовом содержании (которое можно представить также как схему) другого концепта. Рассмотрим пример такой метафоры: пусть некто назван нами «Солнышко!». Солнышко – это метафора. Мы используем не весь объем понятия «солнце» (небесное светило, день и ночь, тепло и т.д.), а лишь ту часть, которая связана с одним из вариантов непосредственного восприятия солнца: яркий утренний свет, свежесть, радость, таинственность. И лишь затем на эту базу или схему проецируем черты конкретного человека.

Существуют и более сложные виды метафор или аналогий. Например, если мы сравним фирму с живым организмом, то далее можем заимствовать такие характеристики организма, как воспроизводство, развитие, безопасность и др. Такая метафора как бы задает нам уже готовый трафарет для моделирования (выполняет роль модели-конфигуратора по терминологии В.А. Лефевра и Г.П. Щедровицкого).

Метафоры и аналогии имеют то же назначение, что и словосочетания и грамматические формы: генерация нового смысла, нового значения, нового знания. Принципиальным во всех этих случаях является тот факт, что в тексте или речи эти новые смыслы, значения и знания присутствуют в сильно сокращенном и «зашифрованном» виде, скорее как ключи, используемые для генерации этих новых знаний, чем как непосредственно сами знания. Для получения же знаний (смыслов, значений) необходима интеллектуальная деятельность субъекта, развертывающая зашифрованное в сообщении знание. Ну, и, в общем-то, как хорошо известно, такая дешифрация может быть осуществлена не единственным образом, и зависит от знаний и предпочтений субъекта.

3. Существует большое количество понятий, содержание которых зависит от развития общества, науки и техники, уровня познания, и вообще от эволюции Вселенной и жизни. В качестве примера рассмотрим влияние эволюции на содержание таких понятий, как интеллект и информация. В эволюции живого можно выделить порядка десятка «больших скачков» [8], на каждом из которых существенно менялись принципы обработки информации, а соответственно менялось само содержание понятий информации и интеллекта: самовоспроизводящиеся молекулярные циклы, вирусы, бактерии, простейшие, насекомые, рептилии, млекопитающие, человек. На рис. 2 дано упрощенное представление об изменении состава (содержания) понятий в процессе эволюции.

При рассмотрении рис. 2 очевидным является тот факт, что простейшие и насекомые, например, не используют языки понятийного уровня, не используют символьных представлений, и вследствие этого не обладают мышлением и сознанием. По этой причине для описания интеллекта на их уровне не требуется использования представлений о понятийном языке, символах и т.п. И наоборот, для описания человеческого интеллекта требуется использование понятийного языка и мышления. Из этих рассуждений следуют два ключевых вывода:

– доказательство наличия интеллекта у какой-либо системы вовсе не доказывает, что этот интеллект сравним с человеческим. Есть задачи принципиально недостижимые для интеллекта более низких уровней. Так, и самолет, и телега являются средством передвижения, но при использовании телеги в принципе невозможно достичь Америки;

– не существует разумного обоснования для существования единого понятия интеллекта (а также информации): понятие интеллекта для нижних уровней эволюции не позволяет объяснить свойства интеллекта высших уровней, и, наоборот, содержание понятия «интеллект» для высших уровней эволюции является существенно избыточным для описания низших уровней эволюции.

Таким образом, само по себе понятие «интеллект» указывает скорее не на конкретное содержание, а на обобщающий функциональный принцип, типа «средства передвижения» (точнее, устройства по обработке информации). После указания функционального принципа, необходимо, вообще говоря, конкретизировать описание. Но мы этого в явном виде не делаем, ограничиваясь неформализуемым (обычно) уточнением в результате учета контекста. Но часто контекста недостаточно для конкретизации, или мы его неправильно интерпретируем.

Вот где источник трудностей, парадоксов и проблем при изучении понятий информации, интеллекта и мышления. Какой же компьютер справится с анализом сообщений, содержащих эти понятия, когда и человек-то зачастую неспособен это адекватно сделать. Выход представляется тот же самый, что использовали эскимосы относительно понятия «снег» (правда, безотносительно к эволюции): для всех понятий, описывающих эволюцию или развитие, необходимо выделить качественно различающиеся этапы, и для каждого из этих этапов создать частные понятия, которые объясняются и описываются через термины, присущие данному этапу эволюции или развития.

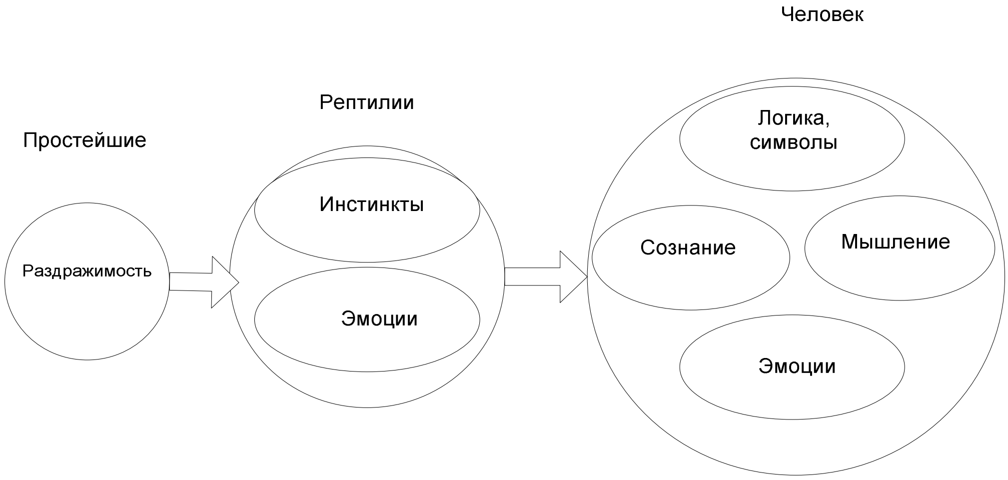


Рисунок 2 – Усложнение понятия интеллекта в процессе эволюции

4) Существует также многозначность смыслов сообщения [1]. Один из источников многозначности – это использование шуток, юмора, намеков, иносказаний, притч и т.п. Второй источник связан с тем, что мы не просто читаем текст (в частности, художественный), а соотносим его со всем личным опытом, переживаниями, стремлениями, идеями и т.п. В результате в процессе чтения мы получаем несколько параллельно развивающихся и взаимосвязанных уровней смысла [1], причем именно высшие уровни являются наиболее важными, индивидуальными и могут быть совершенно различными у разных людей, читающих один и тот же текст. Отметим, что в отличие от компьютера человек легко (и часто с удовольствием) работает с многовариантными, многозначными и неопределенными смыслами, причем, не выбирая одно значение (и отбрасывая остальные), а одновременно «ведет» различные смыслы во время понимания и размышления.

2 Принципы формирования смысла и значения

С формальной стороны суть формирования значения заключается в том, что сообщению каким-то образом ставится в соответствие нечто другое, имеющее отношение к пониманию, интеллекту и мышлению. В связи с этим возникает ряд вопросов:

– что именно ставится в соответствие?

– каковы критерии выбора соответствия (законы перевода системы символов в нечто другое), и откуда, собственно, возникают сами эти критерии?

– как на физическом уровне реализуется установление соответствия?

– для каких минимальных единиц сообщения устанавливается соответствие (посимвольно, между понятиями, между предложениями и т.д.)?

Д. Хофштадтер [9, с. 51] предложил конструктивные соображения по ответу на первый вопрос: «между значением в формальных системах и значением в языке есть важное различие. Различие это заключается в том, что, выучив значение какого-либо слова, мы составляем затем новые предложения, основанные на этом значении. В определенном смысле значение становится *активным*, так как оно порождает новые правила создания предложений… В формальной системе значение должно оставаться *пассивным*; мы можем прочитывать каждую строчку в зависимости от значения символов, ее составляющих, но нам не позволено создавать новые теоремы, основываясь на значениях, которые мы придаем этим символам». Принимая трактовку Хофштадтера, сформулируем ее в более явном виде:

– в формальной системе получение значения заключается в установлении соответствия между двумя системами символов. Сам процесс установления соответствия заключается в том, что из исходной системы символов (сообщения) по определенным заранее правилам получается второй набор символов (значение);

– человек воспринимает исходное сообщение как некий набор значений. Но этот набор не окончательный, он продуцирует некоторые активные процессы, которые приводят к генерации дополнительного и отличного от исходного значения (этот тезис полностью совпадает с результатами данной работы, изложенными в п. 1). Более конкретно вопросы генерации нового значения рассмотрим далее.

Важные соображения по принципам формирования смысла на базе образов и концептов приводятся Ю.Р. Валькманом. Так, в работе [3] он приходит к выводу о целостности образа, невозможности его описания с помощью перечня характеристик: «Сейчас уже понято, что образ далеко не всегда «распродается» на признаки, существенные характеристики. А если и делится на них, то непонятно, по какому основанию, на основе каких критериев». Он также приходит к выводу о невозможности сведения смысла сообщения к сумме составляющих смыслов (что полностью совпадает с выводами, полученными в первой части данной работы, но с помощью аргументации иного типа) [4]: «Под *смыслом* применительно к вербальному образу (тексту) и, в частности, к минимальной единице этого текста понимается *целостное содержание какого-либо высказывания, не сводимое к значениям составляющих его частей и элементов, но само определяющее эти значения*».

Последующий анализ принципов формирования значения, проводимый в данной работе, существенно опирается на результаты, полученные в процитированных работах Д. Хофштадтера и Ю.Р. Валькмана.

Итак, что же именно ставится в соответствие в процессе понимания сообщения? Согласно полученным в п. 1 выводам, для каждого крупного этапа эволюции должно быть разработано свое понимание информации и интеллекта, а, равно как и производных от них понятий, в частности, смысла, значения и др. При этом для всех этапов эволюции, а также для искусственного интеллекта процесс анализа сообщения можно выразить одной общей схемой-последовательностью (рис. 3.):

Содержание приведенной на рис. 3 схемы существенно разнится для разных ситуаций. Кратко рассмотрим наиболее характерные из них.

Для машины Тьюринга сообщением является часть ленты вычисляющего устройства, на которой в двоичном коде записана входная информация [10, с.24-26]. Критерии анализа представляют собой «таблицу команд» или программу, записанную на другой ленте (как устройстве памяти), либо другом участке той же ленты. Анализ сообщения реализуется в исполнительном устройстве путем использования таблицы команд. И результат анализа есть информация, записанная на еще одном участке ленты. Эту ин-

формацию можно также охарактеризовать как интерпретацию или значение сообщения. Характерной особенностью машины Тьюринга является взаимная независимость сообщения от программы его анализа. В то же время результат анализа (значение) определяется не только текстом сообщения, но и используемой для анализа данного сообщения программой. Иначе говоря, значение сообщения в компьютере, вообще говоря, не просто не единственно, а может быть каким угодно. Это и не удивительно, поскольку в самом определении компьютерного значения есть лишь одно ограничивающее требование: наличие соответствия между заданной системой символов (сообщением), и любой другой системой.

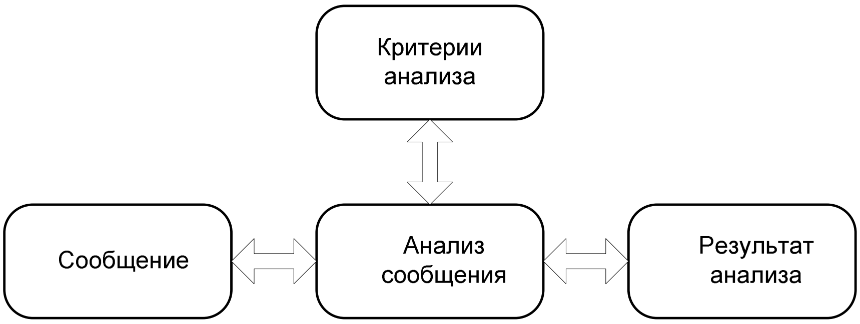


Рисунок 3 – Общая схема анализа сообщения

Рассмотрим другой важный случай: клетка живого организма как инструмент получения значения из сообщения, в какой-то степени аналогичный мозгу и компьютеру. В качестве сообщения служит геном, а в качестве значения или смысла – жизнедеятельность клетки, реализуемая через процессы. К отличительным чертам этого способа обработки информации по сравнению с человеческим мозгом и компьютером относятся следующие:

– не используется символьное и вообще любое носителенезависимое представление информации;

– нет выделенного собственно информационного представления результата анализа сообщения (генома), информация тесно взаимосвязана с механизмами реализации.

Образно говоря, живой организм представляет собой большой набор «читателей» (в человеческом организме несколько десятков триллионов клеток), извлекающих из одного и того же текста (генома) различные значения в зависимости от фенотипа, типа ткани, возраста организма, и, в какой-то мере, внешней среды. В клетке, в отличие от компьютера, сообщение (геном) и систему его анализа уже нельзя считать независимыми. В частности, хорошо известно, что внешние по отношению к геному процессы могут менять «силу» генов, вплоть до их «включения» или «выключения». В связи с этим в сообществе исследователей возникла дискуссия: а где находится информация, предписывающая деятельность клетки: только в геноме или в клетке в целом? Д. Хофштадтер наглядно выразил эту ситуацию [9, с. 192]: «у нас имеются два противоречивых взгляда на природу информации в генотипе. Один из них утверждает, что, поскольку такое большое количество информации содержится вне ДНК, мы должны рассматривать ДНК не более как очень сложный набор пусковых механизмов, что-то вроде кнопок на музыкальном автомате; другой взгляд – что *вся информация* содержится в ДНК, только в очень неявной форме». Представляется, что две крайние позиции, выраженные в цитате, не являются противоречащими, если под неявной формой представления информации понимать ее сжатое, зашифрованное

представление. В этом случае для извлечения зашифрованной информации требуется некоторая дополнительная информация. И чем сильнее сжата информация в сообщении, тем больше требуется внешней информации (эта внешняя информация содержится в устройстве анализа сообщений). Информация может быть понята, получить значение и смысл, только путем интерпретации ее в соответствующем устройстве, предназначенном для обработки именно этого типа информации [11, с. 52].

Получение значения в человеческом мозгу опирается на целый ряд информационных сущностей, отсутствующих при работе генного механизма клетки. Эти информационные сущности описываются такими понятиями, как символ, язык, концепт, гештальт, понимание, мышление, сознание и др. Для полного представления о понимании сообщения требуется также знание нейрофизиологии: каким образом те или иные информационные сущности представлены и обрабатываются на нейрофизиологическом уровне, какими участками мозга, как эти участки мозга взаимодействуют между собой и т.п. По-видимому, такое полное представление о понимании может быть получено в обозримое время, но объем теоретических и экспериментальных работ, которые должны быть проведены для этого, достаточно велик. В данной работе предлагается существенно более простая схема, основное назначение которой – упорядочение представлений о ключевых этапах результатов работы мозга (а не нейрофизиологической основы) в процессе понимания сообщения (рис. 4).

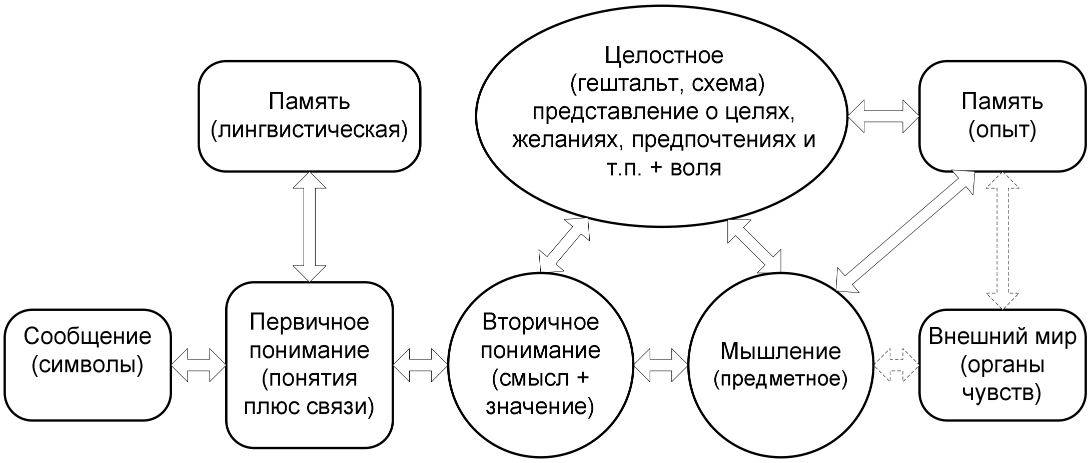


Рисунок 4 – Схема формирования смысла и значения у человека

Человек способен воспринимать информацию не только в непосредственной (ощущения, образы), но и в символьной форме. Суть использования символов заключается в том, что они заменяют собой другие, вообще говоря, значительно более сложные сущности. Таким образом, в процессе восприятия символьного сообщения происходит преобразование относительно простой символьной формы в существенно более сложное когнитивное и нейрофизиологическое (название зависит от того, какой уровень представления мы рассматриваем) представление. Вопрос в том, насколько оно более сложное? В рамках современной компьютерной идеологии обычно ограничиваются уровнем понятий и их качеств и отношений. Данные нейрофизиологии свидетельствуют, что такой (или близкий к этому) уровень понимания символьных сообщений действительно существует. Но это только начало процесса понимания! В когнитивной лингвистике общепринято считать процесс понимания более сложным [1], [2], с этим согласен и ряд специалистов по искусственному интеллекту, в частности [3], [4], [9], а также п.1 данной работы. К слову сказать, существует тяжелое психическое

заболевание, приводящее к умственной неполноценности, называемое аутизмом. Существенным для этого заболевания является именно буквальное понимание слов, и непонимание смысла сообщения в целом [12].

Таким образом, после первого этапа «буквального» (более строго – по-понятийного) понимания идет второй этап, включающий в себя создание смысла и/или значения сообщения. В процессе создания этого более точного уровня смысла возможно появление новых идей, понятий, целей персонажей и др., отсутствующих в исходном сообщении. Часть механизмов создания и представления смысла рассмотрена в п.1. данной работы. Но существует еще один мощный механизм создания нового смысла: избирательное восприятие различных частей сообщения. Избирательность может быть обусловлена целым рядом факторов: различным уровнем доверия к разным источникам (не заслуживающий доверия источник просто не воспринимается, или трактуется с точностью до наоборот), предпочтение к отдельным названиям, идеям и т.д., в результате чего из сообщения «вырезаются» только значимые для субъекта восприятия участки. С точки зрения избирательности содержания, ситуация аналогична «чтению» генома на клеточном уровне. Образно говоря, при таком подходе понимание сообщения больше напоминает игру на флейте (регулировка с помощью клапанов потоков воздуха и воспроизводимых звуков), чем движение по шоссе.

Заметим, что, вообще говоря, существует два основных режима восприятия сообщений: 1) режим «запроса», при котором информация в сообщении, да и сами сообщения, воспринимается избирательно. Для компьютера к этому режиму относятся, в частности, запрос в базы данных или знаний, поиск в Интернет (при этом весь объем информации в Интернет можно рассматривать как одно очень большое сообщение) и др.; 2) режим полного, неизбирательного восприятия. Как аргументируется выше, человек всегда в той или иной мере работает в режиме «запроса». В то же время современная компьютерная лингвистика ориентирована преимущественно (а в большинстве случаев – полностью) на режим полного чтения.

После этого предварительного рассмотрения особенностей интерпретации сообщения в различных когнитивных системах перейдем к рассмотрению вопроса, а что же соответствует в когнитивной системе сообщению, и что есть понимание? Представляется, что это два разных вопроса – соответствие сообщению шире по своему объему процесса понимания. Для описания соответствия сообщения у человека необходимо использовать сразу несколько уровней или слоев:

– соответствие на нейрофизиологическом уровне, т.е., какие именно структуры мозга как физического тела участвуют в отображении содержания сообщения;

– соответствие на процессном уровне, т.е., какие процессы (временная и пространственная структура, взаимодействие, отношения причинности с другими процессами и т.д.) соответствуют содержанию сообщения;

– соответствие на информационном уровне, т.е., что именно в сообщении и каким образом соответствует элементам некоторой обобщенной базы знаний, находящейся в мозгу субъекта восприятия сообщения?

Обычно перечисленные уровни исследуются независимо, в крайнем случае, попарно. Но очевидно, что полностью понять процессы понимания можно лишь путем одновременного изучения всех трех уровней как некоей единой системы.

Последний уровень соответствия (информационное соответствие) можно назвать также смыслом или значением. В лингвистике принято считать смысл внутренним, не формализуемым, или частично формализуемым свойством психики [2]. Выразим ту же мысль в несколько иной трактовке, более пригодной для развития и моделирования.

Смыслу сообщения соответствует все богатство информационной, процессной и нейрофизиологической реализации. Он является «воплощенным» (по англ. embodied), также как мышление и сознание. Смысл обладает (для одного и того же субъекта) более или менее одинаковым содержанием, и может быть формализован лишь частично. Значение – это та часть смысла, которая может быть формализована. Значение всегда уже смысла. Поскольку формализация может быть осуществлена различными способами (полнота, инструмент, цели и т.д.), то любое сообщение может иметь множество значений, даже для одного и того же субъекта. Логическое мышление реализуется на уровне значений. И если логические выводы входят в противоречие со смыслом, то возможно обратное влияние значения на смысл (упрощенно говоря, происходит диалог левого и правого полушарий). В процессе общения мы обмениваемся значениями, пытаясь затем довести их до смысла в процессе понимания, но часто делаем это с заметной степенью неадекватности. Именно об этой принципиальной неполноте общения писал Ф.Н. Тютчев:

Как сердцу высказать себя?
Другому как понять тебя?
Поймет ли он, чем ты живешь?
Мысль изреченная есть ложь…

Понимание следует трактовать как процесс выделения ключевых составляющих смысла сообщения. Как смысл, так и понимание могут изменяться (обогащаться, уточняться и т.д.) в процессе размышлений над сообщением или его частью. Поскольку понимание концентрируется на немногих, ключевых идеях сообщения, то оно легче подвержено формализации, чем смысл. Но в случае неудовлетворительности понимания (по любой причине), уточнение и увеличение адекватности мы можем реализовать лишь «погружаясь» в смысл и вытаскивая из него в качестве ключевых новые, более адекватные представления.

Рассмотрим механизмы выбора при генерации смысла сообщения. Для наглядности используем режим поиска нужной информации человеком (соответствует режиму запроса для компьютера). Характерным для человека является то, что он в состоянии отыскать необходимую информацию, представленную в совершенно различной форме, в частности:

– другими понятиями, эти понятия могут быть синонимами; терминами, описывающими аналогичное явление в других науках; обобщениями данного термина; членами одного и того же ряда и др.;

– информацию, выраженную с помощью тропов (метафор, аналогий и др.);

– выбирать необходимую информацию, разбросанную по разным частям сообщения, или даже по разным сообщениям, и строить на основе этой информации компактное, целостное представление;

– сопоставлять информацию, представленную в виде текста, схем, графиков, формул, рисунков и др.

Объяснить эти способности человека можно, только если предположить у него наличие следующих способностей:

1) Человек легко и свободно преобразовывает информацию, представленную в самой различной форме, к одному и тому же виду;

2) Человек *знает*, что он ищет, и имеет компактную, целостную форму представления этого знания. Мы называем эту целостную форму представления знаний гештальтом, образом или схемой.

В процессе поиска нужной информации человек извлекает из сообщения нужную информацию в виде схемы и сравнивает полученную схему со схемой, возникшей на основе запроса. При удовлетворительной степени сходства схем он работает с извлеченной из текста схемой далее, в противном случае ищет далее.

Для реализации следующего уровня компьютерной семантики необходимо понимать, что *схема принципиально несводима к перечню ее элементов.* Это верно уже на уровне технических устройств. Так, для составления списка элементов радиотехнической схемы требуется совсем немного времени, и нет необходимости в специальном образовании. В то же время для получения понимания того, как работает схема, требуется гораздо больше времени, и соответствующее образование. При этом специалист способен увидеть общность схем, представленных совершенно различными элементами, и с различной топологией связей. Что уж тогда говорить о гораздо более сложных мысленных схемах, к тому же лишь частично формализованных. Из представленных доводов следует вывод, что современные компьютерные методы работы с текстами (в частности, поиск в Интернет), основанные на поиске заданной группы слов, лишь в очень небольшой степени отражают специфику работы с информацией человека, и не могут быть положены в основу прорывных методов компьютерного анализа текстов.

Что же представляет собой схема (гештальт, мысль, целостность) интеллектуального запроса в плане физической реализации и как ее моделировать?

Из субъективного опыта мы знаем, что такого рода схема представляет собой некоторый относительно устойчивый психический процесс. Относительно – потому что в процессе понимания и мышления идея может видоизменяться, обрастать новыми подробностями, трактовками, нюансами и т.п. Устойчивый – потому что в процессе размышлений остается неизменным некоторое ядро запроса или идеи.

Идея, мысль, возникающая у нас, является уникальным, разовым объектом длительностью свыше 100 мсек (время осознания), и основанная на более быстрых процессах (1 – 10 мсек), характерных для работы отдельного нейрона или небольшой группы близко расположенных нейронов. Вследствие того, что мысль как целостность, как глобальный квазиустойчивый процесс основана на более простых процессах, представляется логичным ввести представление о процессах второго и более высоких порядков как способе организации мышления в мозгу.

К ключевым принципам реализации процессов второго порядка необходимо отнести следующие:

– наличие коллективных эффектов в среде процессов первого порядка, и возникновение обратных стабилизирующих связей именно между коллективами процессов первого порядка;

– иерархический характер обработки информации в человеческом мозгу.

Для того чтобы детализировать далее указанные принципы, необходимо определить среду реализации процессов второго порядка.

В качестве первой из сред выступает головной мозг, в котором эти процессы реализуются как «воплощенные». Для их описания требуется детальное знание строения мозга, нейронов, синапсов, процессов, иерархической организации, ритмических волн и многое другое. Представляется, что это сложная задача, и она может быть решена только большим коллективом исследователей в течение значительного времени (до нескольких десятков лет).

Ко второму типу сред следует отнести все системы обработки информации, в которых информация находится в формализованном виде, безразличном к виду

материального носителя, т.е. компьютерные системы обработки информации. По определению модели как упрощенного описания объекта все формальные методы вместе с физическими средствами их реализации, т.е. компьютерами, являются моделями по отношению к мозгу. Поэтому, строго говоря, для компьютеров мы должны говорить не о процессах второго порядка, а о модельном представлении процессов второго порядка, т.е. о модельном представлении работы мозга. Тем не менее, при использовании достаточно адекватных моделей мы в результате работы компьютера должны получить те же результаты, которые получит мозг, если результаты его работы формализовать.

Существует целый ряд интеллектуальных объектов, используемых при высокоуровневом анализе информации: фреймы, классы (в программировании), автоматы по фон Нейману, искусственные нейроны и др. Все эти интеллектуальные объекты могут быть использованы в качестве основы для построения моделей процессов второго порядка. При этом с точки зрения компьютерной реализации они должны представлять собой некие динамические программные объекты, включающие в себя возможности: а) создания новых объектов типа экземпляра класса с учетом входной информации, знаний и целей субъекта; б) включения во вновь созданный объект уже существующих объектов, в частности, базовых описаний концептов, схем, фреймов, графов; в) рассмотрение в интеллектуальных операциях новых или модифицированных объектов как неких целостностей с точки зрения семантики и др.

Выводы

В компьютере семантика текста и речи сводится к сумме семантики отдельных слов и их связей. Человек же в процессе интерпретации и понимания текста создает существенно отличающееся от исходного текста содержание. Существует целый ряд факторов, участвующих в генерации смысла: части речи и грамматические формы; фигуры речи или тропы, в частности, метафора и аналогия; проблемы определения объема понятий, описывающих эволюцию и развитие; многозначность смысла сообщения.

Понятие интеллекта – это не единое понятие, а обобщающий функциональный принцип, обозначающий объект по обработке информации. В связи с этим для понятия «интеллект», так же, как и для других понятий, описывающих эволюцию и развитие, необходимо выделить наиболее крупные этапы эволюции, и содержание понятий определять для каждого этапа эволюции. Естественно, что каждое из такого рода понятий будет развиваться и обогащаться вполне определенным образом.

Объем и содержание понятия конструируются таким образом, чтобы оптимальным образом отображать реальность, включая как внешний мир, так и внутренние переживания.

В сообщении информация находится в неактивном, зашифрованном виде. Для ее активизации необходимо устройство по извлечению и активизации информации (тело клетки, мозг, компьютер и др.). Чем больше своей внутренней информации содержит такое устройство, и чем она более сложно организована, тем больше информации может быть извлечено из сообщения, и тем больше спектр ее возможных значений.

Смыслу сообщения соответствует все богатство информационной, процессной и нейрофизиологической реализации. Он является «воплощенным» (по англ. embodied), также как мышление и сознание. Значение – это та часть смысла, которая может быть формализована.

Человек мыслит посредством целостных психических сущностей, называемых идеями, гештальтами, образами, схемами и т.д. В процессе понимания сообщения из него выделяются схемы, по возможности соответствующие внутренним схемам, основанных на желаниях и запросах.

Мысль как физическая сущность представляет собой процессы второго и более порядков по отношению к нейронам. К ключевым принципам организации таких процессов являются следующие: они основаны на более быстрых процессах, происходящих в отдельных нейронах или небольших их группах; наличие обратных связей между коллективами процессов первого порядка; иерархический характер обработки информации.

Список литературы

1. Богин Г.И. Обретение способности понимать: Введение в герменевтику / Богин Г.И. – Тверь, 2001. – 731 с.
2. Попова З.Д. Когнитивная лингвистика / З.Д. Попова, И.А. Стернин. – М. : АСТ: Восток-Запад – 2007. – 314 с.
3. Валькман Ю.Р. Целостность образов: О моделировании смысла и понимания / Ю.Р. Валькман // Information Technologies & Knowledge – 2012. – V. 6., № 1. – P.14-25.
4. Валькман Ю.Р. О языке образного мышления / Ю.Р. Валькман, Л.Р. Исмагилова // Материалы конф. «Диалог-2004».
5. Сторож В.В. Семантика и онтологии в естественных и искусственных системах / В.В. Сторож // Open Semantic Technologies for Intelligent Systems – OSTIS-2014 – Минск : БГУИР, 2014. – С. 415-422.
6. Лакофф Дж. Метафоры, которыми мы живем / Дж. Лакофф, М. Джонсон – М. : Эдиториал УРСС – 2004. – 256 с.
7. Lakoff G. The invariance hypothesis: is abstract reason based on image-shemas? / G. Lakoff // Cognitive Linguistics – 1990. – V. 1, № 1. – P. 40-71.
8. Сторож В.В. Принципы эволюции интеллекта / В.В. Сторож // Искусственный интеллект – 2007. – № 1. – С. 296-310.
9. Хофштадтер Д.Р. Гедель, Эшер, Бах: Эта бесконечная гирлянда / Хофштадтер Д.Р. – М. : Бахрах-М – 2001. – 752 с.
10. Тьюринг А. Может ли машина мыслить? М.: Физматлит – 1960 . – 112 с.
11. Корогодин В.И. Информация как основа жизни / В.И. Корогодин, В.Л. Корогодина – Дубна : Феникс, 2000. – 208 с.
12. Glezerman T.B. Autism and the Brain: Neurophenomenological Interpretation / Glezerman T.B. – New York : Springer – 2013. – 304 p.

References

1. Bogin G.I. Obretenie sposobnosti ponimat: Vvedenie v germenevtiku – Tver – 2001. – 731 p.
2. Popova Z.D., Sternin I.A. Kognitivnaya lingvistica – M.: AST: Vostok-Zapad – 2007. – 314 p.
3. Valkman Y.R. Zelostnost obrazov: O modelirovanii smisla I ponimaniya // Information Technologies & Knowledge – 2012. – V.6., N 1. – P.14-25.
4. Valkman Y.R. Ismagilova L.R. O yazike obraznogo mishleniya / Proc. of conf "Dialog-2004".
5. Storozh V.V. Semantica I ontologii v estestvennih I iskustvennih sistemah – In Open Semantic Technologies for Intelligent Systems – OSTIS-2014 – Minsk.: BGUIR, 2014. – p.415-422.
6. Lakoff G., Johnsen M. Metaphors we live by – Chicago: The university of Chicago press – 1980. – 193 p.
7. Lakoff G. The invariance hypothesis: is abstract reason based on image-shemas? // Cognitive Linguistics – 1990. – V.1, N 1. – P.40-71.
8. Storozh V.V. Prinzipi evoluzii intellecta // Iscusstvenniy intellect – 2007. – N 1. – P.296-310.
9. Hofstadter D.R. Godel, Escher, Bach: an Eternal Golden Braid – New York.: Basic Books – 1979. – 752 p.
10. Turing A.M. Computing Machinery and Intelligence // Mind – V.49. – P.433-460.
11. Korogodin V.I., Korogodina V.L. Informaziya kak osnova zhizni – Dubna.: Fenix – 2000. – 208 p.
12. Glezerman T.B. Autism and the Brain: Neurophenomenological Interpretation – New York.: Springer – 2013. – 304 p.

RESUME

V.V. Storozh

Principles of Realization of Semantics
in Natural and Artificial Systems

In work it is shown that the person in the processes of interpretation and understanding of the text creates maintenance essentially different from the initial text. Mechanisms of generation of this new maintenance are considered.

The concept of intelligence is not uniform concept, but the generalizing functional principle, designating object on processing of the information.

The sense of the message there correspond all riches information, process and neurophysiological to realization. It is "embodied", also as mind and consciousness. Meaning is that part of sense which can be formalized.

The person minds by means of holistic mental entity, named ideas, gestalt, pattern, schemes. In the processes of understanding of the message from it the schemes whenever possible corresponding to internal schemes, based on desires and inquiries are allocated.

Статья поступила в редакцию 12.05.2014.